Discussion on ‘Of quantiles and expectiles: consistent scoring functions, Choquet representations and forecast rankings’ by Werner Ehm, Tilmann Gneiting, Alexander Jordan and Fabian Krüger

Pinson, Pierre

Published in:
Journal of the Royal Statistical Society, Series B (Statistical Methodology)

Publication date:
2016

Document Version
Peer reviewed version

Comment of Ehm *et al.* (2016) for JRSSB

Authors of paper:	Werner Ehm, Tilmann Gneiting, Alexander Jordan, Fabian Krüger
Title of paper:	Of quantiles and expectiles: consistent scoring functions, Choquet representations and forecast rankings
Author of comment:	Professor Pierre Pinson
Affiliation:	Technical University of Denmark, Kgs. Lyngby, Denmark
Email address:	ppin@dtu.dk

Comment:

The authors ought to be congratulated for an intriguing and exciting contribution to the science of forecasting and more particularly to forecast verification. Some fundamental contributions to this field can be traced back to the 1950s. Even so, some of the key issues in forecasting still appear to be open problems today. I mainly think here of the link between forecast *quality* and *value*. Those were elegantly defined by Murphy (1993), in short, as the objective correspondence of the forecasts with the process observations for the former, and as the increased benefits from integrating such forecasts in decision processes for the latter. Linking these two is definitely not trivial in a general sense, even though for specific example problems and toy models one can nicely illustrate an existing (or non-existing) connection. For forecasters having to deal with a wealth of decision-makers with different decision problems and loss functions, it can never be possible to consider all potential problems and loss functions to assess the value of their forecaststo these decision-makers.

Consequently, the intriguing nature of that contribution lies in the fact the authors investigated and found another rather elegant path to make a better connection between forecast quality and value. Elementary scoring functions and Choquet-type mixture representations allow defining a simple toolbox to assess whether a forecaster (or a set of forecasts) dominates another, under *any* consistent scoring functions. Even though it cannot readily tell whether this forecaster will yield higher value in all potential decision processes, the strength of the authors' result is something that brings us closer to insuring that forecasts seen as having higher quality should eventually yield higher value, whatever their loss function.

Based on this contribution, one is left wondering how practical this result and so-called Murphy diagrams may be in empirical forecast comparisons. The authors mention that empirical dominance is for the case of an elementary scoring curve lies under another one, *for all* θ. How informative really is the case where curves intersect? Besides, this result is given in a univariate setup only, as if when making a decision at time t, the decision-maker would consider a single variable at time $t + k$ only. In many practical applications, decision makers are to jointly account for information from many variables, possibly various lead times, locations, etc. I am therefore wondering whether such result and diagnostic tool could generalize to the case of multivariate setups.

References

Murphy, A.H. (1993) What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weath. Forecast.*, **8**, 281-293.